BIOTECHNOLOGY
SYSTEMS
BRANCH
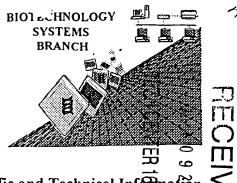
RAW SEQUENCE LISTING
ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer-readable form:

Application Serial Number: *09/435,770 A*

Source: *1652*

Date Processed by STIC: *3 - 30 - 01*

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) **INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,**
2) **TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY**
FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.
PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)
PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 3.0 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

Checker Version 3.0

The Checker Version 3.0 application is a state-of the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 – 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.
Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO).
Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

Checker Version 3.0 can be down loaded from the USPTO website at the following address:
http://www.uspto.gov/web/offices/pac/checker

Raw Sequence Listing Error Summary

ERROR DETECTED	SUGGESTED CORRECTION	SERIAL NUMBER: O9/435,770

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

1 ____ Wrapped Nucleics
The number/text at the end of each line "wrapped" down to the next line.
This may occur if your file was retrieved in a word processor after creating it.
Please adjust your right margin to .3, as this will prevent "wrapping".

2 ____ Wrapped Aminos
The amino acid number/text at the end of each line "wrapped " down to the next line.
This may occur if your file was retrieved in a word processor after creating it.
Please adjust your right margin to .3, as this will prevent "wrapping".

3 ____ Incorrect Line Length
The rules require that a line not exceed 72 characters in length. This includes spaces.

4 ____ Misaligned Amino Acid Numbering
The numbering under each 5th amino acid is misaligned. This may be caused by the use of tabs
between the numbering. It is recommended to delete any tabs and use spacing between the numbers.

5 ____ Non-ASCII
This file was not saved in ASCII (DOS) text, as required by the Sequence Rules.
Please ensure your subsequent submission is saved in ASCII text so that it can be processed.

6 ____ Variable Length
Sequence(s) ____ contain n's or Xaa's which represented more than one residue.
As per the rules, each n or Xaa can only represent a single residue.
Please present the maximum number of each residue having variable length and
indicate in the (ix) feature section that some may be missing.

7 ____ PatentIn ver. 2.0 "bug"
A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid
sequence(s) _____. Normally, PatentIn would automatically generate this section from the
previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section
to the subsequent amino acid sequence. This applies primarily to the mandatory <220>-<223>
sections for Artificial or Unknown sequences.

8 ____ Skipped Sequences (OLD RULES)
Sequence(s) ____ missing. If intentional, please use the following format for each skipped sequence:
(2) INFORMATION FOR SEQ ID NO:X:
(i) SEQUENCE CHARACTERISTICS:(Do not insert any headings under "SEQUENCE CHARACTERISTICS")
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X:
This sequence is intentionally skipped

Please also adjust the "(iii) NUMBER OF SEQUENCES:" response to include the skipped sequence(s).

9 ____ Skipped Sequences (NEW RULES)
Sequence(s) ____ missing. If intentional, please use the following format for each skipped sequence.
<210> sequence id number
€400> sequence id number
000

10 ____ Use of n's or Xaa's (NEW RULES)
Use of n's and/or Xaa's have been detected in the Sequence Listing.
Use of <220> to <223> is MANDATORY if n's or Xaa's are present.
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

11 ____ Use of <213>Organism (NEW RULES)
Sequence(s) _____ are missing this mandatory field or its response.

12 ____ Use of <220>Feature (NEW RULES)
Sequence(s) ____ are missing the <220>Feature and associated headings.
Use of <220> to <223> is MANDATORY if <213>ORGANISM is "Artificial" or "Unknown"
Please explain source of genetic material in <220> to <223> section.
(See "Federal Register," 6/01/98, Vol. 63, No. 104, pp. 29631-32)   (Sec. 1.823 of new Rules)

13 ____ PatentIn ver. 2.0 "bug"
Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted
file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing).
Instead, please use "File Manager" or any other means to copy file to floppy disk.

1652

RAW SEQUENCE LISTING   DATE: 03/30/2001
PATENT APPLICATION: US/09/435,770A   TIME: 10:46:41

Input Set : **A:\YAMAMTO1.txt**
Output Set: **N:\CRF3\03302001\I435770A.raw**

Does Not Comply
Corrected Diskette Needed
See pp. 1-4

```
 4 <110> APPLICANT: YAMAMOTO, Takuo
 5         MARUTA, Kazuhiko
 6         KUBOTA, Michio
 7         FUKUDA, Shigeharu
 8         MIYAKE, Toshio
10 <120> TITLE OF INVENTION: NON-REDUCING SACCHARIDE-FORMING ENZYME,
11        TREHALOSE-RELEASING ENZYME, AND PROCESS FOR PRODUCING
12        SACCHARIDES USING THE ENZYMES
14 <130> FILE REFERENCE: YAMAMOTO=16A
16 <140> CURRENT APPLICATION NUMBER: 09/435,770A
17 <141> CURRENT FILING DATE: 1999-11-08
19 <150> PRIOR APPLICATION NUMBER: JP 258,394/1998
20 <151> PRIOR FILING DATE: 1998-09-11
22 <150> PRIOR APPLICATION NUMBER: JP 352,252/1998
23 <151> PRIOR FILING DATE: 1998-12-11
25 <150> PRIOR APPLICATION NUMBER: JP 16,931/1999
26 <151> PRIOR FILING DATE: 1999-01-26
28 <160> NUMBER OF SEQ ID NOS: 39
30 <170> SOFTWARE: PatentIn Ver. 2.1
```

ERRORED SEQUENCES

```
286 <210> SEQ ID NO: 8
287 <211> LENGTH: 28
288 <212> TYPE: DNA
289 <213> ORGANISM: ARTHROBACTER sp.S34
291 <400> SEQUENCE: 8
```
E--> 292 tttttttaata aaatcaggag gaaaaaat
293 (28)

"Wrapped" nucleic acid number. Move to end of nucleic line.

```
913 <210> SEQ ID NO: (32)
914 <211> LENGTH: 2218
915 <212> TYPE: DNA
916 <213> ORGANISM: Artificial Sequence
918 <220> FEATURE:
919 <223> OTHER INFORMATION: Description of Artificial Sequence:SYNTHETIC
921 <220> FEATURE:
922 <221> NAME/KEY: CDS
923 <222> LOCATION: (477)..(2201)
925 <220> FEATURE:
926 <221> NAME/KEY: 3'UTR
927 <222> LOCATION: (2202)..(2218)
```

Sequence # 32?

E--> 929 <400> SEQUENCE: (33)
```
930 ctgcagctgc tcgcccccgg aacccccgac gtgtaccagg gcacggaacg ctgggaccgg    60
932 tcgctggtgg acccggacaa ccgtcgcccc gtggatttcg ccgcggcatc cgagctgctc   120
934 gaccgcctcg acggcggctg gcggccgccc gtcgacgaga ccggcgcggt caagacgctc   180
936 gtcgtctccc gcgcgctgcg gctgcgccgc gaccggcccg agctgttcac cgcgtaccac   240
```

RAW SEQUENCE LISTING   DATE: 03/30/2001
PATENT APPLICATION: US/09/435,770A   TIME: 10:46:41

Input Set : A:\YAMAMTO1.txt
Output Set: N:\CRF3\03302001\I435770A.raw

```
938 ccggtcacgg cgcgcggcgc gcaggccgag cacctgatcg gcttcgaccg cggcggcgcg   300
940 atcgccctgg ccacccgcct gccgctcggc ctcgccgccg caggcggctg gggcgacacg   360
942 gtcgtcgacg tcggcgagcg gagcctgcgc gacgagctga ccggccgcga ggcccgcgga   420
944 gcggcgcgcg tggccgagtt gttcgccgac taccccgtcg ccctgctggt ggagac atg   479
945                                                             Met
946                                                               1
948 aac cga cga ttc ccg gtc tgg gcg ccc cag gcc gcg cag gtg acg ctc    527
949 Asn Arg Arg Phe Pro Val Trp Ala Pro Gln Ala Ala Gln Val Thr Leu
950             5                   10                  15
952 gtc gtg ggc caa ggc cgc gcc gaa ctc ccg ctg acc cgc gac gag aac    575
953 Val Val Gly Gln Gly Arg Ala Glu Leu Pro Leu Thr Arg Asp Glu Asn
954         20                  25                  30
956 gga tgg tgg gct ctt cag cag ccg tgg gac ggc ggc ccc gac ctc gtc    623
957 Gly Trp Trp Ala Leu Gln Gln Pro Trp Asp Gly Gly Pro Asp Leu Val
958     35                  40                  45
960 gac tac ggc tac ctc gtc gac ggc aag ggc ccc ttc gcc gac ccg cgg    671
961 Asp Tyr Gly Tyr Leu Val Asp Gly Lys Gly Pro Phe Ala Asp Pro Arg
962 50                  55                  60                  65
964 tcg ctg cgg cag ccg cgc ggc gtg cac gag ctc ggc cgc gaa ttc gac    719
965 Ser Leu Arg Gln Pro Arg Gly Val His Glu Leu Gly Arg Glu Phe Asp
966                 70                  75                  80
968 ccc gcc cgc tac gcg tgg ggc gac gac gga tgg cgc ggc cga gac ctc    767
969 Pro Ala Arg Tyr Ala Trp Gly Asp Asp Gly Trp Arg Gly Arg Asp Leu
970             85                  90                  95
972 acc gga gcc gtg atc tac gaa ctg cac gtc ggc acc ttc acc cct gag    815
973 Thr Gly Ala Val Ile Tyr Glu Leu His Val Gly Thr Phe Thr Pro Glu
974         100                 105                 110
976 gga acg ctg gac agc gcc atc cgt cgc ctc gac cac ctg gtg cgc ctc    863
977 Gly Thr Leu Asp Ser Ala Ile Arg Arg Leu Asp His Leu Val Arg Leu
978     115                 120                 125
980 ggc gtc gac gcg gtc gag ctg ctg ccc gtc aac gcg ttc aac ggc acc    911
981 Gly Val Asp Ala Val Glu Leu Leu Pro Val Asn Ala Phe Asn Gly Thr
982 130                 135                 140                 145
984 cac ggc tgg ggc tac gac ggg gtg ctc tgg tac gcg gtg cac gag ccc    959
985 His Gly Trp Gly Tyr Asp Gly Val Leu Trp Tyr Ala Val His Glu Pro
986                 150                 155                 160
988 tac ggc ggc ccg gag gcg tac cag cgc ttc gtc gac gcc tgc cac gcc   1007
989 Tyr Gly Gly Pro Glu Ala Tyr Gln Arg Phe Val Asp Ala Cys His Ala
990             165                 170                 175
992 cgc ggc ctc gcc gtc gtg cag gac gtc gtc tac aac cac ctg ggc ccg   1055
993 Arg Gly Leu Ala Val Val Gln Asp Val Val Tyr Asn His Leu Gly Pro
994         180                 185                 190
996 agc ggc aac cac ctg ccc gac ttc ggc ccc tac ctc ggg tcg ggc gcc   1103
997 Ser Gly Asn His Leu Pro Asp Phe Gly Pro Tyr Leu Gly Ser Gly Ala
998     195                 200                 205
1000 gcc aac acc tgg ggc gac gcg ctg aac ctc gac ggg ccg ctc tcc gac  1151
1001 Ala Asn Thr Trp Gly Asp Ala Leu Asn Leu Asp Gly Pro Leu Ser Asp
1002 210                 215                 220                 225
1004 gag gtg cgg cgg tac atc atc gac aac gcg gtg tac tgg ctg cgc gac  1199
```

RAW SEQUENCE LISTING   DATE: 03/30/2001
PATENT APPLICATION: US/09/435,770A   TIME: 10:46:41

Input Set : A:\YAMAMTO1.txt
Output Set: N:\CRF3\03302001\I435770A.raw

```
1005 Glu Val Arg Arg Tyr Ile Ile Asp Asn Ala Val Tyr Trp Leu Arg Asp
1006             230             235                 240
1008 atg cac gcc gac ggg ctg cgg ctc gac gcc gtg cac gcg ctg cgc gac  1247
1009 Met His Ala Asp Gly Leu Arg Leu Asp Ala Val His Ala Leu Arg Asp
1010             245             250                 255
1012 gcc cgc gcg ctg cac ctg ctc gaa gag ctc gcc gcc cgc gtc gac gag  1295
1013 Ala Arg Ala Leu His Leu Leu Glu Glu Leu Ala Ala Arg Val Asp Glu
1014             260             265                 270
1016 ctg gcg ggc gag ctc ggc cgg ccg ctg acg ctc atc gcc gag agc gac  1343
1017 Leu Ala Gly Glu Leu Gly Arg Pro Leu Thr Leu Ile Ala Glu Ser Asp
1018     275             280                 285
1020 ctg aac gac ccg aag ctg atc cgc tcc cgc gcg gcg cac ggc tac ggc  1391
1021 Leu Asn Asp Pro Lys Leu Ile Arg Ser Arg Ala Ala His Gly Tyr Gly
1022 290             295                 300             305
1024 ctc gac gcc cag tgg gac gac gac gtg cac cac gcg gtg cac gcc aac  1439
1025 Leu Asp Ala Gln Trp Asp Asp Asp Val His His Ala Val His Ala Asn
1026             310             315                 320
1028 gtg acc ggc gag acc gtc ggc tac tac gcc gac ttc ggc ggg ctc ggc  1487
1029 Val Thr Gly Glu Thr Val Gly Tyr Tyr Ala Asp Phe Gly Gly Leu Gly
1030             325             330                 335
1032 gcc ctc gtc aag gtg ttc cag cgc ggc tgg ttc cac gac ggc acc tgg  1535
1033 Ala Leu Val Lys Val Phe Gln Arg Gly Trp Phe His Asp Gly Thr Trp
1034             340             345                 350
1036 tcg agc ttc cgc gag cgg cac cac ggc cgg ccg ctc gac ccc gac atc  1583
1037 Ser Ser Phe Arg Glu Arg His His Gly Arg Pro Leu Asp Pro Asp Ile
1038             355             360                 365
1040 ccg ttc cgc cgg ctc gtc gcc ttc gcg cag gat cac gac cag gtc ggc  1631
1041 Pro Phe Arg Arg Leu Val Ala Phe Ala Gln Asp His Asp Gln Val Gly
1042 370             375                 380             385
1044 aac cga gcg gtc ggc gac cgc atg tcg gcg cag gtc ggc gag ggt tcg  1679
1045 Asn Arg Ala Val Gly Asp Arg Met Ser Ala Gln Val Gly Glu Gly Ser
1046             390             395                 400
1048 ctc gcc gcc gcg gcg gcg ctc gtg ctg ctc ggc ccg ttc acc ccg atg  1727
1049 Leu Ala Ala Ala Ala Ala Leu Val Leu Leu Gly Pro Phe Thr Pro Met
1050             405             410                 415
1052 ctg ttc atg ggc gag gag tgg ggc gcg cgc acc ccg tgg cag ttc ttc  1775
1053 Leu Phe Met Gly Glu Glu Trp Gly Ala Arg Thr Pro Trp Gln Phe Phe
1054             420             425                 430
1056 acc tcc cac ccc gag ccc gag ctg ggg gag gcg acg gcg cgc ggg cgc  1823
1057 Thr Ser His Pro Glu Pro Glu Leu Gly Glu Ala Thr Ala Arg Gly Arg
1058     435             440                 445
1060 atc gcc gag ttc gcc cgc atg ggc tgg gac ccg gca gtc gtg ccc gac  1871
1061 Ile Ala Glu Phe Ala Arg Met Gly Trp Asp Pro Ala Val Val Pro Asp
1062 450             455                 460             465
1064 ccg cag gac ccg gcc acc ttc gcc cgc tcg cac ctg gac tgg tcc gag  1919
1065 Pro Gln Asp Pro Ala Thr Phe Ala Arg Ser His Leu Asp Trp Ser Glu
1066             470             475               · 480
1068 ccc gag cgg gaa ccg cac gcg ggc ctg ctc gcc ttc tac acc gac ctg  1967
1069 Pro Glu Arg Glu Pro His Ala Gly Leu Leu Ala Phe Tyr Thr Asp Leu
```

RAW SEQUENCE LISTING   DATE: 03/30/2001
PATENT APPLICATION: US/09/435,770A   TIME: 10:46:41

Input Set : A:\YAMAMTO1.txt
Output Set: N:\CRF3\03302001\I435770A.raw

```
1070              485              490              495
1072 atc gcg ctg cgg cgc gag ctg ccg gtc gat gcg ccg gcg cgc gag gtg   2015
1073 Ile Ala Leu Arg Arg Glu Leu Pro Val Asp Ala Pro Ala Arg Glu Val
1074         500              505              510
1076 gat gcc gac gag gcg cgc ggc gtc ttc gcg ttc agc cgc ggc ccg ctg   2063
1077 Asp Ala Asp Glu Ala Arg Gly Val Phe Ala Phe Ser Arg Gly Pro Leu
1078    515              520              525
1080 cgg gtc acg gtc gcg ctg cgc ccc gga ccg gtc ggg gtg ccc gag cac   2111
1081 Arg Val Thr Val Ala Leu Arg Pro Gly Pro Val Gly Val Pro Glu His
1082 530              535              540              545
1084 ggg ggc ctc gtg ctc gcc tac ggc gag gtg cgc gcc ggc gcc gcc gga   2159
1085 Gly Gly Leu Val Leu Ala Tyr Gly Glu Val Arg Ala Gly Ala Ala Gly
1086              550              555              560
1088 ctg cac ctc gac ggg ccg gga gcc gcg atc gtg cgc ctc gag           2201
1089 Leu His Leu Asp Gly Pro Gly Ala Ala Ile Val Arg Leu Glu
1090              565              570              575
1092 tgacgcggct gggtacc                                                2218
1166 <210> SEQ ID NO: 39
1167 <211> LENGTH: 22
1168 <212> TYPE: DNA
1169 <213> ORGANISM: Artificial Sequence
1171 <220> FEATURE:
1172 <223> OTHER INFORMATION: Description of Artificial Sequence:SYNTHETIC
1174 <400> SEQUENCE: 39
1175 aggtggttgt agacgacgtc ct                                             22
E--> 1181 20
E--> 1187 15
```

Delete extraneous numerals from the end of the file. They are causing an invalid base count for sequence #39.

VERIFICATION SUMMARY
PATENT APPLICATION: **US/09/435,770A**

DATE: 03/30/2001
TIME: 10:46:42

Input Set : **A:\YAMAMTO1.txt**
Output Set: **N:\CRF3\03302001\I435770A.raw**

L:292 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:28 SEQ:8
L:909 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:31
L:909 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:31
L:909 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:31
L:929 M:212 E: (34) Invalid or duplicate Sequence ID Number, SEQUENCE ID NOS:32 differs:33
L:1181 M:254 E: No. of Bases conflict, LENGTH:Input:20 Counted:22 SEQ:39
M:254 Repeated in SeqNo=39